

Distributed Control of Positive Systems

Anders Rantzer

Abstract—It is shown that several classical open problems in linear system theory, such as computation of stabilizing static output feedback gains and optimization of decentralized controllers, can be readily solved for positive systems using linear programming. In particular, optimal solutions can be verified for large-scale systems using computations that scale linearly with the number of interconnections. The restriction to positive state space models is natural in many applications, including stochastic systems, transportation networks and economics. Moreover, the main results are generalized to frequency domain representations using the notion of positively dominated system.

I. INTRODUCTION

Classical methods for multi-variable control, such as LQG and H_∞ -optimization, suffer from a lack of scalability that make them hard to use for large-scale systems. The difficulties are partly due to computational complexity, but also absence of distributed structure in the resulting controllers. Complexity growth can be traced back to the fact that stability verification of a linear system with n states generally requires a Lyapunov function involving n^2 quadratic terms, even if the system matrices are sparse. In this paper we will see that the situation improves drastically if we restrict attention to closed loop dynamics described by system matrices with nonnegative off-diagonal entries. Then stability and performance can be verified using a Lyapunov function with only n linear terms. Sparsity can be exploited in performance verification and even synthesis of distributed controllers can be done with a complexity that grows linearly with the number of nonzero entries in the system matrices. These observations have far-reaching implications for control engineering:

- 1) The conditions that enable scalable solutions hold naturally in many important application areas, such as stochastic systems, economics, transportation networks, chemical reactions, power systems and ecology.
- 2) The essential mathematical property can be extended to frequency domain models. A sufficient condition is that the transfer functions involved are “positively dominated”.
- 3) In control applications, the assumption of positive dominance need not hold for the open loop process. Instead, a large-scale control system can often be structured into local control loops that give positive dominance, thus enabling scalable methods for optimization of the global performance.

The paper is structured as follows: Sections II-III introduce background literature and notation. Stability criteria for positive systems are cited in section IV. These results are not new, but stated on a form convenient for later use and explained

with emphasis on scalability. Section V extends the stability results to input-output performance. The analysis results are then exploited in section VI for synthesis of stabilizing and optimal controllers using distributed linear programming. Section VII extends the techniques to positively dominated transfer functions. Section VIII explains how Lyapunov inequalities for positive systems can be verified using methods that scale linearly with the number of states and interconnections. Similar methods are used in section IX to prove a general version of the Kalman-Yakubovich-Popov Lemma for positive systems.

II. BACKGROUND

The study of matrices with nonnegative coefficients has a long history, dating back to the Perron-Frobenius Theorem in 1912. A classic book on the topic is [3]. The theory is used in Leontief economics [17], where the states denote nonnegative quantities of commodities. Systems defined by nonnegative matrices (so called positive systems) appear in the study of Markov chains [26], where the states denote nonnegative probabilities and in compartment models [11], where the states could denote populations of species. A nice introduction to the subject is given in [18].

A fundamental property of linear maps described by a positive matrix is that they are contractive in Hilbert’s projective metric [4], [14]. This metric is closely related to the Lyapunov function $\max\{x_1, \dots, x_n\} - \min\{x_1, \dots, x_n\}$, used for example in analysis of consensus algorithms [26], [29]. For more recent contributions, see [21], [27].

A nonlinear counterpart to positive systems is monotone systems, characterized by the property that a partial ordering of initial states is preserved by the dynamics. Such dynamical systems were studied in a series of papers by Hirsch [8], [9], showing that monotonicity generally implies convergence almost everywhere. Positive systems have also gained increasing attention in the control literature during the last decade. See for example [30], [7], [12]. Feedback stabilization of positive linear systems was studied in [16], [23], [22] and basic control theory for nonlinear monotone systems was developed in [1]. A recent result by Tanaka and Langbort [28] shows that decentralized controllers can be optimized for positive systems using semi-definite programming. The criterion is the closed loop H_∞ norm and the authors show that diagonal quadratic storage functions can be used without conservatism. Some of the main results in this paper (first presented in [24]), can be viewed as extensions of that work.

III. NOTATION

Let \mathbb{R}_+ denote the set of nonnegative real numbers. The inequality $X > 0$ ($X \geq 0$) means that all elements of the matrix (or vector) X are positive (nonnegative). For a symmetric

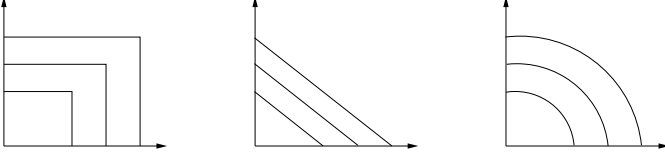


Fig. 1. Level curves of Lyapunov functions corresponding to the conditions (1.2), (1.3) and (1.4) in Proposition 1: If $A\xi < 0$, then $V(x) = \max_i(x_i/\xi_i)$ is a Lyapunov function with rectangular level curves. If $z^T A < 0$, then $V(x) = z^T x$ is a linear Lyapunov function. Finally if $A^T P + PA \prec 0$ and $P \succ 0$, then $V(x) = x^T P x$ is a quadratic Lyapunov function for the system $\dot{x} = Ax$.

matrix X , the inequality $X \succ 0$ means that the matrix is positive definite. The matrix $A \in \mathbb{R}^{n \times n}$ is said to be *Hurwitz* if all eigenvalues have negative real part. It is *Schur* if all eigenvalues are strictly inside the unit circle. Finally, the matrix is said to be *Metzler* if all off-diagonal elements are nonnegative. The notation $\mathbb{H}_{\infty}^{n \times m}$ represents the set of $n \times m$ matrices whose entries are analytic functions in the right half of the complex plane.

IV. DISTRIBUTED STABILITY VERIFICATION

Proposition 1: Given a Metzler matrix $A \in \mathbb{R}^{n \times n}$, the following statements are equivalent:

- (1.1) The matrix A is Hurwitz.
- (1.2) There exists a $\xi \in \mathbb{R}^n$ such that $\xi > 0$ and $A\xi < 0$.
- (1.3) There exists a $z \in \mathbb{R}^n$ such that $z > 0$ and $z^T A < 0$.
- (1.4) There exists a *diagonal* matrix $P \succ 0$ such that $A^T P + PA \prec 0$.
- (1.5) The matrix $-A^{-1}$ exists and has nonnegative entries.

Moreover, if $\xi = (\xi_1, \dots, \xi_n)$ and $z = (z_1, \dots, z_n)$ satisfy the conditions of (1.2) and (1.3) respectively, then $P = \text{diag}(z_1/\xi_1, \dots, z_n/\xi_n)$ satisfies the conditions of (1.4).

Remark 1. Each of the conditions (1.2), (1.3) and (1.4) corresponds to a Lyapunov function of a specific form. See Figure 1.

Remark 2. One of the main observations of this paper is that verification and synthesis of positive control systems can be done with methods that scale linearly with the number of interconnections. For stability, this claim follows directly from Proposition 1: Given ξ , verification of the inequality $A\xi < 0$ requires a number of scalar additions and multiplications that is directly proportional to the number of nonzero elements in the matrix A . In fact, the search for a feasible ξ also scales linearly, since integration of the differential equation $\dot{\xi} = A\xi$ with $\xi(0) = \xi_0$ for an arbitrary $\xi_0 > 0$ generates a feasible $\xi(t)$ in finite time provided that A is Metzler and Hurwitz.

Proof of Proposition 1. The equivalence between (1.1), (1.2), (1.4) and (1.5) is the equivalence between the statements G_{20} , I_{27} , H_{24} and N_{38} in [3, Theorem 6.2.3]. The equivalence between (1.1) and (1.3) is obtained by applying the equivalence between (1.1) and (1.2) to the transpose of A . Moreover, if $\xi = (\xi_1, \dots, \xi_n)$ and $z = (z_1, \dots, z_n)$ satisfy the conditions of (1.2) and (1.3) respectively, then $P = \text{diag}(z_1/\xi_1, \dots, z_n/\xi_n)$ gives $(A^T P + PA)\xi = A^T z + PA\xi < 0$ so the symmetric matrix $A^T P + PA$ is Hurwitz and (1.4) follows. \square

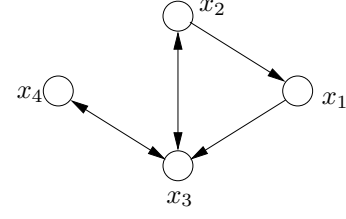


Fig. 2. A graph of interconnected systems. In Example 1 the interpretation is a transportation network and each arrow indicates a transportation link. In Example 2 the interpretation is instead a vehicle formation and each arrow indicates the use of a distance measurement.

Example 1. Linear transportation network. Consider a dynamical system interconnected according to the graph illustrated in Figure 2:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} -1 - \ell_{31} & \ell_{12} & 0 & 0 \\ 0 & 2 - \ell_{12} - \ell_{32} & \ell_{23} & 0 \\ \ell_{31} & \ell_{32} & 3 - \ell_{23} - \ell_{43} & \ell_{34} \\ 0 & 0 & \ell_{43} & -4 - \ell_{34} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \quad (1)$$

The model could for example be used to describe an transportation network connecting four buffers. The states x_1, x_2, x_3, x_4 represent the contents of the buffers and the parameter ℓ_{ij} determines the rate of transfer from buffer j to buffer i . Without such transfer the content of the second and third buffer would grow exponentially due to the diagonal elements 2 and 3, corresponding to unstable internal dynamics of those buffers.

Notice that the dynamics can be written as $\dot{x} = Ax$ where A is a Metzler matrix provided that every ℓ_{ij} is nonnegative. Hence, by Proposition 1, stability is equivalent to existence of numbers $\xi_1, \dots, \xi_4 > 0$ such that

$$\begin{bmatrix} -1 - \ell_{31} & \ell_{12} & 0 & 0 \\ 0 & 2 - \ell_{12} - \ell_{32} & \ell_{23} & 0 \\ \ell_{31} & \ell_{32} & 3 - \ell_{23} - \ell_{43} & \ell_{34} \\ 0 & 0 & \ell_{43} & -4 - \ell_{34} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix} < \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Given these numbers, stability can be verified by a distributed test where the first buffer verifies the first inequality, the second buffer verifies the second and so on. In particular, the relevant test for each buffer only involves parameter values at the local node and the neighboring nodes, so a global model is not needed anywhere. \square

Example 2. Vehicle formation (or distributed Kalman filter). Another system structure, which can be viewed as a dual of the previous one, is the following:

$$\begin{cases} \dot{x}_1 = -x_1 + \ell_{13}(x_3 - x_1) \\ \dot{x}_2 = \ell_{21}(x_1 - x_2) + \ell_{23}(x_3 - x_2) \\ \dot{x}_3 = \ell_{32}(x_2 - x_3) + \ell_{34}(x_4 - x_3) \\ \dot{x}_4 = -4x_4 + \ell_{43}(x_3 - x_4) \end{cases} \quad (2)$$

This model could for example be used to describe a platoon of four vehicles. The parameters ℓ_{ij} represent position adjustments based on distance measurements between the vehicles. The terms $-x_1$ and $-4x_4$ reflect that the first and fourth vehicle are equipped to maintain stable positions on their own, but the second and third vehicle rely on the distance measurements for stabilization. Again, stability can be verified by a distributed test where the first vehicle verifies the first

inequality, the second vehicle verifies the second inequality and so on. \square

A discrete time counterpart to Proposition 1 can be stated as follows:

Proposition 2: For $B \in \mathbb{R}_+^{n \times n}$, the following statements are equivalent:

- (2.1) The matrix B is Schur stable.
- (2.2) There is a $\xi \in \mathbb{R}^n$ such that $\xi > 0$ and $B\xi < \xi$.
- (2.3) There exists a $z \in \mathbb{R}^n$ such that $z > 0$ and $B^T z < z$.
- (2.4) There is a diagonal $P \succ 0$ such that $B^T P B \prec P$.
- (2.5) $(I - B)^{-1}$ exists and has nonnegative entries.

Moreover, if $\xi = (\xi_1, \dots, \xi_n)$ and $z = (z_1, \dots, z_n)$ satisfy the conditions of (2.2) and (2.3) respectively, then $P = \text{diag}(z_1/\xi_1, \dots, z_n/\xi_n)$ satisfies the conditions of (2.4).

Proof. The equivalence between (2.1) and (2.5) is proved by [3, Lemma 6.2.1]. Setting $A = B - I$ gives the equivalence between (2.2), (2.3) and (2.5) from the equivalence between (1.2), (1.3) and (1.5).

Suppose $\xi = (\xi_1, \dots, \xi_n)$ and $z = (z_1, \dots, z_n)$ satisfy the conditions of (2.2) and (2.3) respectively. Set $P = \text{diag}(z_1/\xi_1, \dots, z_n/\xi_n)$ and $y_k = \sqrt{\xi_k} z_k$ for $k = 1, \dots, n$. Then

$$\begin{aligned} P^{-1/2} B^T P B P^{-1/2} y &= P^{-1/2} B^T P B \xi < P^{-1/2} B^T P \xi \\ &= P^{-1/2} B^T z < P^{-1/2} z = y \end{aligned}$$

so $P^{-1/2} B^T P B P^{-1/2}$ is Schur, $P^{-1/2} B^T P B P^{-1/2} \prec I$ and (2.4) follows. Finally, (2.4) shows that $x^T P x$ is a positive definite Lyapunov function for the system $x^+ = Bx$, so (2.1) follows from (2.4). \square

V. INPUT-OUTPUT PERFORMANCE OF POSITIVE SYSTEMS

We will now move beyond stability and discuss input-output performance. A connection between stability and performance is established by the following theorem:

Theorem 3: Suppose that $\mathbf{G}(s) = C(sI - A)^{-1}B + D$ with $A \in \mathbb{R}^{n \times n}$ Metzler and $B \in \mathbb{R}_+^{n \times 1}$, $C \in \mathbb{R}_+^{1 \times n}$, $D \in \mathbb{R}_+$. Define $\|\mathbf{G}\|_\infty = \sup_\omega |G(i\omega)|$. Then the following statements are equivalent:

- (3.1) The matrix A is Hurwitz and $\|\mathbf{G}\|_\infty < \gamma$.
- (3.2) The matrix $\begin{bmatrix} A & B \\ C & D - \gamma \end{bmatrix}$ is Hurwitz.
- (3.3) There exists a diagonal $P \succ 0$ such that all solutions to $\dot{x} = Ax + Bw$ satisfy

$$\frac{d}{dt} x(t)^T P x(t) + |Cx(t) + Dw(t)|^2 \leq \gamma^2 |w(t)|^2 \quad (3)$$

with strict inequality when $(x(t), w(t)) \neq (0, 0)$.

- (3.4) There exists $p \in \mathbb{R}^n$ such that $p > 0$ and all solutions to $\dot{x} = Ax + Bw$ satisfy

$$\frac{d}{dt} (p^T |x(t)|) + |Cx(t) + Dw(t)| \leq \gamma |w(t)| \quad (4)$$

with strict inequality when $(x(t), w(t)) \neq (0, 0)$.

Remark 3. The equivalence between (3.1) and (3.3) was earlier proved in [28].

Remark 4. An consequence of Theorem 3 is that the four equivalent conditions not only guarantee that γ is a gain bound on $\mathbf{G}(s)$, but on every operator of the form $C(I - \Delta A)^{-1} \Delta B + D$ where $\Delta(x) = (\Delta_1(x_1), \dots, \Delta_n(x_n))$ and every Δ_k is a passive scalar operator. In fact, (3) can be written as

$$2x(t)^T P(Ax(t) + Bw(t)) + |Cx(t) + Dw(t)|^2 \leq \gamma^2 |w(t)|^2$$

Integrating the inequality over $t \in [0, \tau]$ and using the passivity conditions $\int_0^\tau x_k(Ax_k + Bw_k)dt \geq 0$ gives the desired gain bound $\int_0^\tau |Cx(t) + Dw(t)|^2 dt \leq \gamma^2 \int_0^\tau |w(t)|^2 dt$.

Proof of Theorem 3 First note that the maximum $\max_\omega |\mathbf{G}(i\omega)|$ must be attained at $\omega = 0$ since

$$\begin{aligned} |\mathbf{G}(i\omega)| &= \left| D + \int_0^\infty e^{-i\omega t} C e^{At} B dt \right| \\ &\leq D + \int_0^\infty C e^{At} B dt = D - CA^{-1}B = \mathbf{G}(0) \end{aligned}$$

Hence $\|\mathbf{G}\|_\infty < \gamma$ may equivalently be written

$$D - CA^{-1}B < \gamma$$

Suppose that (3.1) holds. Then $D - CA^{-1}B < \gamma$. By Proposition 1 there exists $\xi > 0$ such that $A\xi < 0$. Define $x = \xi - A^{-1}B$. Then $x > 0$ since $-A^{-1} \geq 0$. Moreover

$$Ax + B = A\xi < 0$$

If ξ is sufficiently small, we also get $Cx + D < \gamma$ so

$$\begin{bmatrix} A & B \\ C & D - \gamma \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} < \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (5)$$

and (3.2) holds. Conversely, (3.2) implies (5) for some $x > 0$, so A is Hurwitz and

$$-A^{-1}B < x \quad D - CA^{-1}B < Cx + D < \gamma$$

Hence (3.1) follows and equivalence between (3.1) and (3.2) has been established.

Next assume that (3.3) holds. For $w = 0$, the inequality (3) implies $\frac{d}{dt} x(t)^T P x(t) \leq 0$ with strict inequality for $x(t) \neq 0$, so A must be Hurwitz. Integrating (3) along solutions to $\dot{x} = Ax + Bw$ with $x(0) = 0$ gives

$$\int_0^T |Cx(t) + Dw(t)|^2 dt \leq \gamma^2 \int_0^T |w(t)|^2 dt$$

so $\|\mathbf{G}\|_\infty \leq \gamma$ and (3.1) follows. The opposite implication, from (3.1) to (3.3), is known from [28] and is also a consequence of Theorem 14, to be proved in the last section of this paper: Condition (3.1) implies (14.1) of Theorem 14 with strict inequality for

$$Q = \begin{bmatrix} C^T C & C^T D \\ DC & D^2 - \gamma^2 \end{bmatrix}$$

Hence (14.3) of Theorem 14 implies existence of a diagonal $P \succ 0$ such that

$$2x^T P(Ax + Bw) + |Cx + Dw|^2 \leq \gamma^2 |w|^2$$

with strict inequality for $(x, w) \neq (0, 0)$. Condition (3.3) follows and the equivalence between (3.1) and (3.3) is established.

Assuming that (3.4) holds, consider $x(t) > 0$, $w(t) > 0$ and rewrite (4) as

$$p^T(Ax + Bw) + Cx + Dw < \gamma w$$

Hence

$$\begin{bmatrix} A & B \\ C & D - \gamma \end{bmatrix}^T \begin{bmatrix} p \\ 1 \end{bmatrix} < 0 \quad (6)$$

and (3.2) follows by Proposition 1.

Finally, assuming (3.2), we get (6) for some $p > 0$, so

$$\begin{aligned} & \frac{d}{dt} (p^T |x(t)|) + |Cx(t) + Dw(t)| \\ & \leq p^T \left| \frac{dx}{dt}(t) \right| + C|x(t)| + D|w(t)| \\ & \leq p^T (A|x(t)| + B|w(t)|) + C|x(t)| + D|w(t)| \leq 0 \end{aligned}$$

with strict inequality when $(x(t), w(t)) \neq (0, 0)$. This proves (3.4) and the proof is complete. \square

A discrete time version can be stated as follows:

Theorem 4: Let $\mathbf{G}(z) = C(zI - A)^{-1}B + D$ where $A \in \mathbb{R}_+^{n \times n}$, $B \in \mathbb{R}_+^{n \times 1}$, $C \in \mathbb{R}_+^{1 \times n}$ and $D \in \mathbb{R}_+$. Define $\|\mathbf{G}\|_\infty = \sup_\omega |\mathbf{G}(e^{i\omega})|$. Then the following conditions are equivalent:

(4.1) The matrix A is Schur and $\|\mathbf{G}\|_\infty < \gamma$.

(4.2) The matrix $\begin{bmatrix} A & B \\ \gamma^{-1}C & \gamma^{-1}D \end{bmatrix}$ is Schur.

(4.3) There exists a diagonal matrix $P \succ 0$ such that

$$|x(t+1)|_P^2 + |Cx(t) + Dw(t)|^2 \leq |x(t)|_P^2 + \gamma^2 |w(t)|^2$$

for all solutions to $x(t+1) = Ax(t) + Bw(t)$.

(4.4) There exists a $p \in \mathbb{R}^n$ such that $p > 0$ and

$$p^T |x(t+1)| + |Cx(t) + Dw(t)| \leq p^T |x(t)| + \gamma |w(t)|$$

for all solutions to $x(t+1) = Ax(t) + Bw(t)$.

A proof of Theorem 4 is given in the appendix. We conclude this section by pointing out that for scalar positive systems, all induced norms are equal:

Theorem 5: For a scalar impulse response $g(t)$ and $w \in \mathbf{L}_p[0, \infty)$, let $g*w$ denote the convolution of g and w . Suppose that $g(t) \geq 0$ and $\int_0^\infty g(t)dt < \infty$. Then the induced norm $\|g\|_{p\text{-ind}} = \sup_w \frac{\|g*w\|_p}{\|w\|_p}$ satisfies

$$\|g\|_{p\text{-ind}} = \int_0^\infty g(t)dt \quad p \in [1, \infty]$$

Proof. It is well known that $\|g\|_{2\text{-ind}} = \max_\omega |\mathbf{G}(e^{i\omega})|$ where $\mathbf{G}(s) = \int_0^\infty g(t)e^{-st}dt$. When $g(t) \geq 0$, the maximum must be attained at $\omega = 0$ since

$$|\mathbf{G}(i\omega)| = \left| \int_0^\infty g(t)e^{-i\omega t}dt \right| \leq \int_0^\infty g(t)dt = \mathbf{G}(0)$$

Moreover

$$\begin{aligned} \|y(t)\|_1 &= \int_0^\infty \left| \int_0^t g(t-\tau)w(\tau)d\tau \right| dt \\ &\leq \int_0^\infty \int_0^t g(t-\tau)|w(\tau)|d\tau dt \\ &= \int_0^\infty \left(\int_\tau^\infty g(t-\tau)dt \right) |w(\tau)|d\tau \\ &= \left(\int_0^\infty g(t)dt \right) \|w\|_1 \end{aligned}$$

with equality when $w(t) \geq 0$ for all t . Similarly

$$\begin{aligned} |y(t)| &= \left| \int_0^\infty g(\tau)w(t-\tau)d\tau \right| \\ &\leq \int_0^\infty g(\tau)|w(t-\tau)|d\tau \\ &\leq \left(\int_0^\infty g(\tau)d\tau \right) \|w\|_\infty \end{aligned}$$

with equality if w is constant. Hence the desired equality

$$\|g\|_{p\text{-ind}} = \int_0^\infty g(t)dt \quad (7)$$

has been proved for $p = 1$, $p = 2$ and $p = \infty$. The Riesz-Thorin convexity theorem [10, Theorem 7.1.12] shows that $\|g\|_{p\text{-ind}}$ is a convex function of p for $1 \leq p \leq \infty$, so (7) must hold for all $p \in [1, \infty]$. \square

VI. DISTRIBUTED CONTROL SYNTHESIS BY LINEAR PROGRAMMING

Equipped with scalable analysis methods for stability and performance, we are now ready to consider synthesis of controllers by distributed optimization. We will start by revisiting an example of section IV.

Example 3 Consider again the transportation network (1), this time with the flow parameters $\ell_{31} = 2$, $\ell_{34} = 1$ and $\ell_{43} = 2$ fixed:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} -3 & \ell_{12} & 0 & 0 \\ 0 & 2 - \ell_{12} - \ell_{32} & \ell_{23} & 0 \\ 2 & \ell_{32} & 1 - \ell_{23} & 1 \\ 0 & 0 & 2 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \quad (8)$$

We will ask the question how to find the remaining parameters ℓ_{12} , ℓ_{23} and ℓ_{32} in the interval $[0, 10]$ such that the closed loop system (8) becomes stable. According to Proposition 1, stability is equivalent to existence of $\xi_1, \dots, \xi_4 > 0$ such that

$$\begin{bmatrix} -3 & \ell_{12} & 0 & 0 \\ 0 & 2 - \ell_{12} - \ell_{32} & \ell_{23} & 0 \\ 2 & \ell_{32} & 1 - \ell_{23} & 1 \\ 0 & 0 & 2 & -5 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix} < 0$$

At first sight, this looks like a difficult problem due to multiplications between the two categories of parameters. However, a closer look suggests the introduction of new variables: $\mu_{12} := \ell_{12}\xi_2$, $\mu_{32} := \ell_{32}\xi_2$ and $\mu_{23} := \ell_{23}\xi_3$. The problem

then reduces to linear programming: Find $\xi_1, \xi_2, \xi_3, \xi_4 > 0$ and $\mu_{12}, \mu_{32}, \mu_{23} \geq 0$ such that

$$\begin{bmatrix} -3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 2 & 0 & 1 & 1 \\ 0 & 0 & 2 & -5 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mu_{12} \\ \mu_{32} \\ \mu_{23} \end{bmatrix} < 0$$

$$\mu_{12} \leq 10\xi_2 \quad \mu_{32} \leq 10\xi_2 \quad \mu_{23} \leq 10\xi_3$$

with the solution $(\xi_1, \xi_2, \xi_3, \xi_4) = (43, 12.8, 10.1, 4.2)$ and $(\mu_{12}, \mu_{32}, \mu_{23}) = (128, 0, 101)$. The corresponding stabilizing gains can then be computed as

$$\ell_{12} = \mu_{12}/\xi_2 = 10 \quad \ell_{32} = \mu_{32}/\xi_2 = 0 \quad \ell_{23} = \mu_{23}/\xi_3 = 10$$

□

The idea can be generalized into the following theorem:

Theorem 6: Let the matrices $A \in \mathbb{R}^{n \times n}$, $E \in \mathbb{R}^{n \times m}$, $F \in \mathbb{R}^{m \times n}$, $K \in \mathbb{R}^{m \times m}$ be given and let \mathcal{D} be the set of $m \times m$ diagonal matrices with entries in $[0, 1]$. Suppose that $(I - LK)^{-1}$ exists and $A + E(I - LK)^{-1}LF$ is Metzler for all $L \in \mathcal{D}$. If F and K have nonnegative coefficients, then the following two conditions are equivalent:

- (6.1) There exists $L \in \mathcal{D}$ such that $A + E(I - LK)^{-1}LF$ is Hurwitz.
- (6.2) There exist $\xi \in \mathbb{R}_+^n$, $\mu \in \mathbb{R}_+^m$ with $\mu \leq F\xi + K\mu$ and $A\xi + E\mu < 0$.

Alternatively, if E and K have nonnegative coefficients, then (6.1) is equivalent to

- (6.3) There exist $p \in \mathbb{R}_+^n$, $q \in \mathbb{R}_+^m$ with $q \leq E^T p + K^T q$ and $A^T p + F^T q < 0$.

Remark 5. If the diagonal elements of \mathcal{D} are restricted to \mathbb{R}_+ instead of $[0, 1]$, then the condition $\mu \leq F\xi + K\mu$ is replaced by $0 \leq F\xi + K\mu$.

Remark 6. Each row of the vector inequalities can be verified separately to get a distributed test.

Remark 7. It is natural to compare the expression $A + E(I - LK)^{-1}LF$ with the “state feedback” expression $A + BL$ of standard linear quadratic optimal control. A major difference is the presence of F and K which make the optimization into a problem of “static output feedback” rather than state feedback. Another difference is the diagonally structured L instead of a full matrix. The diagonal structure gives a much higher degree of flexibility, particularly in the specification of distributed controllers.

Proof of Theorem 6. Suppose (6.1) holds. Let $A + E(I - LK)^{-1}LF$ be Hurwitz and define $\xi \in \mathbb{R}_+^n$ with $[A + E(I - LK)^{-1}LF]\xi < 0$. Let $\mu = (I - LK)^{-1}LF\xi$. Then $\mu = L(F\xi + K\mu)$ and $A\xi + E\mu = (A + E(I - LK)^{-1}LF)\xi < 0$.

Conversely, suppose that (6.2) holds. Choose $L \in \mathcal{D}$ to get $\mu = (I - LK)^{-1}LF\xi$. Then

$$[A + E(I - LK)^{-1}LF]\xi = A\xi + E\mu < 0$$

so $A + E(I - LK)^{-1}LF$ is Hurwitz. The equivalence between (6.1) and (6.3) follows immediately by replacing $A + E(I - LK)^{-1}LF$ with its transpose. □

Combining Theorem 3 with Theorem 6 also gives a linear programming formulation of the problem to minimize input-output gain:

Corollary 7: Let \mathcal{D} be the set of $m \times m$ diagonal matrices with entries in $[0, 1]$. Suppose that D is scalar and that $A + ELF$ is Metzler for all $L \in \mathcal{D}$.

If the matrices B, C, D and F have nonnegative coefficients, then the following two conditions are equivalent:

- (7.1) There exists $L \in \mathcal{D}$ such that $A + ELF$ is Hurwitz and

$$\|C[sI - (A + ELF)]^{-1}B + D\|_\infty < \gamma. \quad (9)$$

- (7.2) There exist $\xi \in \mathbb{R}_+^n$, $\mu \in \mathbb{R}_+^m$ with

$$A\xi + E\mu + B < 0 \quad C\xi + D < \gamma \quad \mu \leq F\xi$$

If ξ, μ satisfy (7.2), then (7.1) holds for every L such that $\mu = LF\xi$.

Alternatively, if B, C, D and E are nonnegative, then (7.1) is equivalent to

- (7.3) There exist $p \in \mathbb{R}_+^n$, $q \in \mathbb{R}_+^m$ with

$$A^T p + F^T q + C^T < 0 \quad B^T p + D < \gamma \quad q \leq E^T p$$

If p, q satisfy (7.3), then (7.1) holds for every L such that $q = LE^T p$.

Remark 8. It is interesting to compare Corollary 7 with the corresponding synthesis method proposed in [28], which exploits (1.4) of Proposition 1 as criterion for stability rather than (1.2) and (1.3). An advantage of the approach in [28] is that the Metzler property of the closed loop system matrix can be enforced in the synthesis procedure as a constraint, rather than being verified a priori for all $L \in \mathcal{D}$. On the other hand, the linear programming approach proposed here has a simpler structure, where the distributed and scalable nature of the conditions is apparent. Tanaka and Langbort do not discuss the scalability aspects of their approach in [28], but a scalable decomposition of their conditions will be developed in Theorem 12 of this paper.

Proof. According to Theorem 3, condition (7.1) holds if and only if there exists $\xi \in \mathbb{R}_+^n$ with

$$\begin{bmatrix} A + ELF & B \\ C & D - \gamma \end{bmatrix} \begin{bmatrix} \xi \\ 1 \end{bmatrix} < 0 \quad (10)$$

Given (10), the inequalities of (7.2) hold with $\mu = LF\xi$. Conversely, given (7.2), the inequalities of (10) follow provided that $\mu = LF\xi$. This proves the desired equivalence between (7.1) and (7.2). The equivalence between (7.1) and (7.3) follows immediately by replacing $G(s)$ with its transpose. □

Example 4. Disturbance rejection in vehicle formation. Consider the vehicle formation model

$$\begin{cases} \dot{x}_1 = -x_1 + \ell_{13}(x_3 - x_1) + w_1 \\ \dot{x}_2 = \ell_{21}(x_1 - x_2) + \ell_{23}(x_3 - x_2) + w_2 \\ \dot{x}_3 = \ell_{32}(x_2 - x_3) + \ell_{34}(x_4 - x_3) + w_3 \\ \dot{x}_4 = -4x_4 + \ell_{43}(x_3 - x_4) + w_4 \end{cases} \quad (11)$$

where w_1, \dots, w_4 are external disturbances acting on the vehicles. Our problem is to find feedback gains $\ell_{ij} \in [0, 1]$

that stabilize the formation and minimize the gain from w to x . The problem can be solved by applying linear programming to condition (7.2) with

$$\begin{aligned} A &= \text{diag}\{-1, 0, 0, -4\} & D &= 0 \\ C &= \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} & K &= 0 \\ E &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} & B &= \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \\ L &= \text{diag}\{\ell_{13}, \ell_{21}, \ell_{23}, \ell_{32}, \ell_{34}, \ell_{43}\} \\ F &= \begin{pmatrix} -1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \end{aligned}$$

The optimal $L = \text{diag}\{0, 1, 1, 0, 1, 0\}$ gives $\gamma = 4.125$. If we instead put $B = \begin{pmatrix} 10 & 10 & 1 & 1 \end{pmatrix}^T$, the minimal value $\gamma = 15.562$ is attained with $L = \text{diag}\{1, 1, 1, 0, 1, 0\}$. Conversely, $B = \begin{pmatrix} 1 & 1 & 10 & 10 \end{pmatrix}^T$ gives the minimal value $\gamma = 12.750$ for $L = \text{diag}\{0, 1, 0, 1, 1, 0\}$. \square

A discrete time counterpart to Theorem 6 is given without proof:

Theorem 8: Let \mathcal{D} be the set of $m \times m$ diagonal matrices with entries in $[0, 1]$. Given $A \in \mathbb{R}^{n \times n}$, $E \in \mathbb{R}^{n \times m}$, $F \in \mathbb{R}^{m \times n}$, $K \in \mathbb{R}^{m \times m}$, suppose that $(I - LK)^{-1}$ exists and $A + E(I - LK)^{-1}LF$ is non-negative for all $L \in \mathcal{D}$. If F and K have nonnegative coefficients, then the following are equivalent:

- (8.1) $A + E(I - LK)^{-1}LF$ is Schur for some $L \in \mathcal{D}$.
- (8.2) There exist $\xi \in \mathbb{R}_+^n$, $\mu \in \mathbb{R}_+^m$ with $\mu \leq F\xi + K\mu$ and $A\xi + E\mu < \xi$.

Alternatively, if E and K have nonnegative coefficients, then (8.1) is equivalent to

- (8.3) There exist $p \in \mathbb{R}_+^n$, $q \in \mathbb{R}_+^m$ with $q \leq E^T p + K^T q$ and $A^T p + F^T q < p$.

VII. POSITIVELY DOMINATED SYSTEMS

So far, the emphasis has been on state space models. However, for many applications input-output models are more natural as a starting point. In this section, we will therefore extend the main ideas of the previous sections to such models. First we need to define a notion of positivity for input-output models. One option would be to work with non-negative impulse responses like in Theorem 5. However, to verify for a given rational transfer function that the impulse response is non-negative has proved to be NP-hard! See [5] for the discrete time problem and [2] for continuous time. Instead we will use the following definition.

A stable rational transfer matrix \mathbf{G} is called *positively dominated* if every matrix entry satisfies $|\mathbf{G}_{jk}(i\omega)| \leq \mathbf{G}_{jk}(0)$ for all $\omega \in \mathbb{R}$. (More generally the term is used for $\mathbf{G} \in \mathbb{H}_\infty^{m \times n}$

if $\|\mathbf{G}_{jk}\|_\infty = \liminf_{\epsilon \rightarrow 0} \mathbf{G}_{jk}(\epsilon)$). The set of all such matrices is denoted $\mathbb{DH}_\infty^{m \times n}$. The essential frequency inequality can be tested by semi-definite programming, since

$$\left| \frac{b(i\omega)}{a(i\omega)} \right| \leq \frac{b(0)}{a(0)} \quad \text{for } \omega \in \mathbb{R}$$

if and only if the polynomial $|a(i\omega)|^2 b(0)^2 - |b(i\omega)|^2 a(0)^2$ is a sum of squares.

Some properties of positively dominated transfer functions follow immediately:

Proposition 9: Let $\mathbf{G}, \mathbf{H} \in \mathbb{DH}_\infty^{n \times n}$. Then $\mathbf{GH} \in \mathbb{DH}_\infty^{n \times n}$ and $a\mathbf{G} + b\mathbf{H} \in \mathbb{DH}_\infty^{n \times n}$ when $a, b \in \mathbb{R}_+$. Moreover $\|\mathbf{G}\|_\infty = \|\mathbf{G}(0)\|$.

The following property is also fundamental:

Theorem 10: Let $\mathbf{G} \in \mathbb{DH}_\infty^{n \times n}$. Then $(I - \mathbf{G})^{-1} \in \mathbb{DH}_\infty^{n \times n}$ if and only if $\mathbf{G}(0)$ is Schur.

Proof. That $(I - \mathbf{G})^{-1}$ is stable and positively dominated implies that $[I - \mathbf{G}(0)]^{-1}$ exists and is nonnegative, so $\mathbf{G}(0)$ must be Schur according to Proposition 2. On the other hand, if $\mathbf{G}(0)$ is Schur we may choose $\xi \in \mathbb{R}_+$ and $\epsilon > 0$ with $\mathbf{G}(0)\xi < (1 - \epsilon)\xi$. Then for every $z \in \mathbb{C}^n$ with $0 < |z| < \xi$ and $s \in \mathbb{C}$ with $\text{Re } s \geq 0$ we have

$$|\mathbf{G}(s)^t z| \leq \mathbf{G}(0)^t |z| < (1 - \epsilon)^t |z| \quad \text{for } t = 1, 2, 3, \dots$$

Hence $\sum_{k=0}^\infty \mathbf{G}(s)^k z$ is convergent and bounded above by $\sum_{k=0}^\infty \mathbf{G}(0)^k |z| = [I - \mathbf{G}(0)]^{-1} |z|$. The sum of the series solves the equation $[I - \mathbf{G}(s)] \sum_{k=0}^\infty \mathbf{G}(s)^k z = z$, so therefore $\sum_{k=0}^\infty \mathbf{G}(s)^k z = [I - \mathbf{G}(s)]^{-1} z$. This proves $(I - \mathbf{G})^{-1}$ is stable and positively dominated and the proof is complete. \square

Corollary 7 can now be generalized as follows:

Theorem 11: Let \mathcal{D} be the set of $m \times m$ diagonal matrices with entries in $[0, 1]$. Suppose that $\mathbf{B} \in \mathbb{DH}_\infty^{n \times 1}$, $\mathbf{C} \in \mathbb{DH}_\infty^{1 \times n}$, $\mathbf{D} \in \mathbb{DH}_\infty$ and $\mathbf{A} + \mathbf{E}\mathbf{L}\mathbf{F} \in \mathbb{DH}_\infty^{n \times n}$ for all $L \in \mathcal{D}$.

If $\mathbf{F} \in \mathbb{DH}_\infty^{m \times n}$, then the following two conditions are equivalent:

- (11.1) There is $L \in \mathcal{D}$ with $(I - \mathbf{A} - \mathbf{E}\mathbf{L}\mathbf{F})^{-1} \in \mathbb{DH}_\infty^{n \times n}$ and $\|\mathbf{C}(I - \mathbf{A} - \mathbf{E}\mathbf{L}\mathbf{F})^{-1}\mathbf{B} + \mathbf{D}\|_\infty < \gamma$.
- (11.2) There exist $\xi \in \mathbb{R}_+^n$, $\mu \in \mathbb{R}_+^m$ with

$$\begin{aligned} \mathbf{A}(0)\xi + \mathbf{E}(0)\mu + \mathbf{B}(0) &< \xi \\ \mathbf{C}(0)\xi + \mathbf{D}(0) &< \gamma \\ \mathbf{F}(0)\xi &\geq \mu \end{aligned}$$

If ξ, μ satisfy (11.2), then (11.1) holds for every L such that $\mu = \mathbf{L}\mathbf{F}(0)\xi$.

Alternatively, if $\mathbf{E} \in \mathbb{DH}_\infty^{n \times m}$, then (11.1) is equivalent to

- (11.3) There exist $p \in \mathbb{R}_+^n$, $q \in \mathbb{R}_+^m$ with

$$\begin{aligned} \mathbf{A}(0)^T p + \mathbf{F}(0)^T q + \mathbf{C}(0)^T &< p \\ \mathbf{B}(0)^T p + \mathbf{D}(0) &< \gamma \\ \mathbf{E}(0)^T p &\geq q \end{aligned}$$

If p, q satisfy (11.3), then (11.1) holds for every L such that $q = \mathbf{L}\mathbf{E}(0)^T p$.

Proof. Theorem 10 shows that (11.1) holds if and only if $\mathbf{A}(0) - \mathbf{E}(0)\mathbf{L}\mathbf{F}(0)$ is Schur and $\mathbf{C}[I - \mathbf{A}(0) - \mathbf{E}(0)\mathbf{L}\mathbf{F}(0)]^{-1}\mathbf{B}(0) + \mathbf{D}(0) < \gamma$. According to Theorem 4, this is true if and only if

$$\begin{bmatrix} \mathbf{A}(0) + \mathbf{E}(0)\mathbf{L}\mathbf{F}(0) & \mathbf{B}(0) \\ \gamma^{-1}\mathbf{C}(0) & \gamma^{-1}\mathbf{D}(0) \end{bmatrix} \quad (12)$$

is Schur. By Proposition 2 this is equivalent to existence of $\xi \in \mathbb{R}_+^n$ such that

$$\begin{bmatrix} \mathbf{A}(0) + \mathbf{E}(0)\mathbf{L}\mathbf{F}(0) & \mathbf{B}(0) \\ \gamma^{-1}\mathbf{C}(0) & \gamma^{-1}\mathbf{D}(0) \end{bmatrix} \begin{bmatrix} \xi \\ 1 \end{bmatrix} < \begin{bmatrix} \xi \\ 1 \end{bmatrix}$$

This is equivalent to (11.2) if we set $\mu = \mathbf{L}\mathbf{F}(0)\xi$, so the desired equivalence between (11.1) and (11.2) in Theorem 11 follows. The equivalence between (11.1) and (11.3) is obtained by replacing $G(s)$ with its transpose. \square

Example 5. Formation of vehicles with inertia. In Example 3, the inputs and disturbances were supposed to have an immediate impact on the vehicle velocities, i.e. the inertia of the vehicles was neglected. Alternatively, a model that takes the inertia into account can be stated as follows:

$$\ddot{x}_i = \sum_j \ell_{ij}(x_j - x_i) + u_i + w_i \quad i = 1, \dots, N$$

where u_i is an external control force, w_i is a disturbance and ℓ_{ij} is the spring constant between the vehicles i and j . Suppose that local control laws $u_i = -k_i x_i - d_i \dot{x}_i$ are given and consider the problem to find spring constants $\ell_{ij} \in [0, \bar{\ell}_{ij}]$ that minimize the gain from w_1 to x_1 .

The closed loop system has the following frequency domain description

$$\begin{aligned} & \left(s^2 + d_i s + k_i + \sum_j \bar{\ell}_{ij} \right) X_i(s) \\ &= \sum_j \left(\ell_{ij} X_j(s) + (\bar{\ell}_{ij} - \ell_{ij}) X_i(s) \right) + W_i(s) \end{aligned}$$

Similarly to Example 3, we write this on matrix form as

$$X = (\mathbf{A} + \mathbf{E}\mathbf{L}\mathbf{F})X + \mathbf{B}W$$

The transfer matrices \mathbf{B} , \mathbf{E} and $\mathbf{A} + \mathbf{E}\mathbf{L}\mathbf{F}$ are positively dominated for all $L \in \mathcal{D}$ provided that $d_i \geq k_i + \sum_j \bar{\ell}_{ij}$. Hence Theorem 11 can then be applied to find the optimal spring constants. Notice that ℓ_{ij} and ℓ_{ji} must be optimized separately, even though by symmetry they must be equal at optimum. \square

VIII. SCALABLE VERIFICATION OF THE LYAPUNOV INEQUALITY

In the preceding sections we have derived scalable conditions for verification of stability and optimality based on the linear inequalities in (1.2) and (1.3) of Proposition 1. The condition (1.4) is also useful for applications, but it less obvious how to verify the Lyapunov inequality $A^T P + P A \prec 0$ in a test that scales linearly with the number of non-zero entries of A . However, the following result can be used for this purpose:

Theorem 12: A symmetric Metzler matrix with $2m$ non-zero off-diagonal entries is negative semi-definite if and only if it can be written as a sum of m negative semi-definite matrices where each term has only four non-zero entries.

The proof of Theorem 12 will be based on the following minor modification of [13, Theorem 3.1]:

Proposition 13 (Positive Quadratic Programming):

Suppose M_0, \dots, M_K are Metzler and $b_1, \dots, b_K \in \mathbb{R}$. Then

$$\begin{aligned} \max_{x \in \mathbb{R}_+^n} \quad & x^T M_0 x &= \max_{X \succeq 0} \quad & \text{trace}(M_0 X) \\ \text{s.t.} \quad & x^T M_k x \geq b_k && \text{trace}(M_k X) \geq b_k \\ & k = 1, \dots, K && k = 1, \dots, K \end{aligned} \quad (13)$$

The value on the right hand side remains the same if the condition $X \succeq 0$ is relaxed to $X \in \mathbb{X}$, where \mathbb{X} is the set of symmetric matrices $(x_{ij}) \in \mathbb{R}^{n \times n}$ satisfying $x_{ii} \geq 0$ and $x_{ij}^2 \leq x_{ii}x_{jj}$ for all i, j . Moreover, if there exists a matrix X in the interior of \mathbb{X} with $\text{trace}(M_k X) \geq b_k$ for every k , then the maximum of (13) is equal to the minimum of $-\sum_k \tau_k b_k$ over $\tau_1, \dots, \tau_K \geq 0$ such that $M_0 + \sum_k \tau_k M_k \preceq 0$.

Remark 9. The problem on the right of (13) is always convex and readily solvable by semidefinite programming. The problem on the left is generally not a convex program, since the matrices M_k may be indefinite. However, the maximization on the left is concave in (x_1^2, \dots, x_n^2) [19]. This is because every product $x_i x_j$ is the geometric mean of two such variables, hence concave [6, p. 74].

Remark 10. The second statement of Proposition 13 is important for scalability, since the condition $X \in \mathbb{X}$ has a natural decomposition and only entries of X that correspond to non-zero entries of M_k need to be taken into account.

Proof of Proposition 13. Every x satisfying the constraints on the left hand side of (13) corresponds to a matrix $X = x x^T$ satisfying the constraints on the right hand side. This shows that the right hand side of (13) is at least as big as the left.

On the other hand, let $X = (x_{ij})$ be a positive definite matrix. In particular, the diagonal elements x_{11}, \dots, x_{nn} are non-negative and $x_{ij} \leq \sqrt{x_{ii}x_{jj}}$. Let $x = (\sqrt{x_{11}}, \dots, \sqrt{x_{nn}})$. Then the matrix $x x^T$ has the same diagonal elements as X , but has off-diagonal elements $\sqrt{x_{ii}x_{jj}}$ instead of x_{ij} . The fact that $x x^T$ has off-diagonal elements at least as big as those of X , together with the assumption that the matrices M_k are Metzler, gives $x^T M_k x \geq \text{trace}(M_k X)$ for $k = 1, \dots, K$. This shows that the left hand side of (13) is at least as big as the right. Nothing changes if X is not positive definite but $X \in \mathbb{X}$, so the second statement is also proved.

For the last statement, note that the conditions $\text{trace}(M_k X) \geq b_k$ are linear in X , so strong duality holds [25, Theorem 28.2] and the right hand side of (13) has a finite maximum if and only if $M_0 + \sum_{k=1}^K \tau_k M_k \preceq 0$ for some $\tau_1, \dots, \tau_K \geq 0$. \square

Proof of Theorem 12. Let \mathcal{E} be the set of indices (k, l) of non-zero off-diagonal entries in M . Define

$$\mathbb{X}_{\mathcal{E}} = \{X \in \mathbb{R}^{n \times n} : [e_k \ e_l]^T X [e_k \ e_l] \geq 0 \text{ for all } (k, l) \in \mathcal{E}\}$$

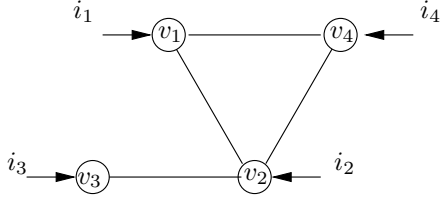


Fig. 3. Illustration of the power transmission network described in Example 5 and 6. Example 5 describes a purely resistive network under static conditions, while Example 6 treats a dynamic network with inductive transmission lines and capacitances at the nodes.

where e_1, \dots, e_n are the unit vectors in \mathbb{R}^n . If M is negative semi-definite, then

$$\begin{aligned} 0 &= \max_{|x| \leq 1} x^T M x = \max_{X \in \mathbb{X}_{\mathcal{E}}} \text{trace}(MX) \\ &= \min_{N_{kl} \geq 0} \max_{X \in \mathbb{R}^{n \times n}} \text{trace}(MX) + \sum_{(k,l) \in \mathcal{E}} \text{trace} \left(N_{kl} [e_k \ e_l]^T X [e_k \ e_l] \right) \\ &= \min_{N_{kl} \geq 0} \max_{X \in \mathbb{R}^{n \times n}} \text{trace} \left[\left(M + \sum_{(k,l) \in \mathcal{E}} [e_k \ e_l] N_{kl} [e_k \ e_l]^T \right) X \right] \end{aligned}$$

where $N_{kl} \in \mathbb{R}^{2 \times 2}$ for every k and l . In particular, there exists a choice of the matrices N_{kl} that makes $M + \sum_{(k,l) \in \mathcal{E}} [e_k \ e_l] N_{kl} [e_k \ e_l]^T = 0$. This completes the proof. \square

To emphasize that Proposition 13 is of independent application interest, we include an example from [15]:

Example 6. Optimal power flow in DC network. Consider a DC power transmission network as in Figure 3. All nodes are subject to constraints of the form $i_k v_k \leq p_k$. For generating nodes p_k represents the generator capacity. For power consuming loads i_k and p_k are negative and $-p_k$ represents the power demand.

Every connection has a known admittance $y_{jk} = y_{kj} \geq 0$. In particular, the current flowing from node 1 to node 2 equals $y_{12}(v_1 - v_2)$. Writing Kirchhoff's current law for all nodes in Figure 3 gives

$$\begin{cases} i_1 = y_{12}(v_1 - v_2) + y_{14}(v_1 - v_4) \\ i_2 = y_{12}(v_2 - v_1) + y_{23}(v_2 - v_3) + y_{24}(v_2 - v_4) \\ i_3 = y_{23}(v_3 - v_2) \\ i_4 = y_{14}(v_4 - v_1) + y_{24}(v_4 - v_2) \end{cases} \quad (14)$$

Suppose every link has a capacity bound c_{ij} on the transferred power and every node has upper and lower bounds on the voltage according to $v_k^{\min} \leq v_k \leq v_k^{\max}$. Then the problem to minimize the power losses in the network subject to constraints on power demands, voltage and link capacities can be written

$$\begin{aligned} &\text{Minimize} && i_1 v_1 + \dots + i_N v_N \\ &\text{subject to} && (14) \text{ with } v_k i_k \leq p_k \\ & && v_k^{\min} \leq v_k \leq v_k^{\max} \\ & && y_{jk}(v_k - v_j)^2 \leq c_{jk} \\ & && \text{for } j, k = 1, \dots, N \end{aligned}$$

This is a quadratic optimization problem of the form handled by Proposition 13. \square

IX. THE KYP LEMMA FOR POSITIVE SYSTEMS

Input-output gain is certainly not the only way to quantify the performance of a linear time-invariant system. A more general class of specifications known as Integral Quadratic Constraints [20] can be tested using the Kalman-Yakubovich-Popov lemma. It is therefore of interest to see that the corresponding result of [28] for positive systems can be generalized the following way:

Theorem 14: Let $A \in \mathbb{R}^{n \times n}$ be Metzler and Hurwitz, while $B \in \mathbb{R}_+^{n \times m}$ and the pair (A, B) is controllable. Suppose that all entries of $Q \in \mathbb{R}^{(n+m) \times (n+m)}$ are nonnegative, except for the last m diagonal elements. Then the following statements are equivalent:

(14.1) For $\omega \in [0, \infty]$ is is true that

$$\begin{bmatrix} (i\omega I - A)^{-1} B \\ I \end{bmatrix}^* Q \begin{bmatrix} (i\omega I - A)^{-1} B \\ I \end{bmatrix} \preceq 0$$

$$(14.2) \quad \begin{bmatrix} -A^{-1} B \\ I \end{bmatrix}^* Q \begin{bmatrix} -A^{-1} B \\ I \end{bmatrix} \preceq 0.$$

(14.3) There exists a diagonal $P \succeq 0$ such that

$$Q + \begin{bmatrix} A^T P + P A & P B \\ B^T P & 0 \end{bmatrix} \preceq 0$$

Moreover, if all inequalities are taken to be strict, then the equivalences hold even without the controllability assumption.

Remark 11. For $A = -1$, $B = 0$, $Q = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, condition (14.1) holds, but not (14.3). This demonstrates that the controllability assumption is essential.

Proof. One at a time, we will prove the implications (14.1) \Rightarrow (14.2) \Rightarrow (14.3) \Rightarrow (14.1). Putting $\omega = 0$ immediately gives (14.2) from (14.1).

Assume that (14.2) holds. The matrix $-A^{-1}$ is nonnegative, so $\begin{bmatrix} x \\ w \end{bmatrix}^T Q \begin{bmatrix} x \\ w \end{bmatrix} \leq 0$ for all $x \in \mathbb{R}_+^n$, $w \in \mathbb{R}_+^m$ with

$$x \leq -A^{-1} B w \quad (15)$$

The inequality (15) follows (by multiplication with $-A^{-1}$ from the left) from the constraint $0 \leq Ax + Bw$, which can also be written $0 \leq A_i x + B_i w$ for $i = 1, \dots, n$, where A_i and B_i denote the i :th rows of A and B respectively. For non-negative x and w , this is equivalent to

$$0 \leq x_i (A_i x + B_i w) \quad i = 1, \dots, n \quad (16)$$

Hence (14.2) implies that $\begin{bmatrix} x \\ w \end{bmatrix}^T Q \begin{bmatrix} x \\ w \end{bmatrix} \leq 0$ for $x \in \mathbb{R}_+^n$, $w \in \mathbb{R}_+^m$ satisfying (16). Proposition 13 will next be used to verify existence of $\tau_1, \dots, \tau_n \geq 0$ such that the quadratic form

$$\sigma(x, w) = \begin{bmatrix} x \\ w \end{bmatrix}^T Q \begin{bmatrix} x \\ w \end{bmatrix} + \sum_i \tau_i x_i (A_i x + B_i w)$$

is negative semi-definite. However, the application of Proposition 13 requires existence of a positive definite X such that all diagonal elements of

$$\begin{bmatrix} A & B \end{bmatrix} X \begin{bmatrix} I \\ 0 \end{bmatrix}$$

are positive. The pair (A, B) is controllable, so there exists K that make all eigenvalues of $A + BK$ unstable and therefore $(A + BK)Z + Z(A + BK)^T = I$ has a symmetric positive definite solution Z . Hence the desired X can be constructed as

$$X = \begin{bmatrix} Z & ZK^T \\ KZ & * \end{bmatrix}$$

where the lower right corner is chosen big enough to make $X \succ 0$.

Define $P = \text{diag}(\tau_1, \dots, \tau_n) \succeq 0$. Then σ being negative definite means that

$$Q + \begin{bmatrix} A^T P + PA & PB \\ B^T P & 0 \end{bmatrix} \preceq 0$$

so (14.3) follows.

Finally, assume that (14.3) holds. Integrating $\sigma(x(t), w(t))$ over time gives

$$0 \geq \int_0^\infty \left(\begin{bmatrix} x \\ w \end{bmatrix}^T Q \begin{bmatrix} x \\ w \end{bmatrix} + x^T P(Ax + Bw) \right) dt$$

For square integrable solutions to $\dot{x} = Ax + Bw$, $x(0) = 0$ we get

$$\begin{aligned} 0 &\geq \int_0^\infty \left(\begin{bmatrix} x \\ w \end{bmatrix}^T Q \begin{bmatrix} x \\ w \end{bmatrix} + \frac{d}{dt}(x^T P x / 2) \right) dt \\ &= \int_0^\infty \begin{bmatrix} x(t) \\ w(t) \end{bmatrix}^T Q \begin{bmatrix} x(t) \\ w(t) \end{bmatrix} dt \end{aligned}$$

which in frequency domain implies (14.1). Hence (14.1) \Rightarrow (14.2) \Rightarrow (14.3) \Rightarrow (14.1).

For strict inequalities, the proofs that (14.3) \Rightarrow (14.1) \Rightarrow (14.2) remain the same. Assuming that (14.2) holds with strict inequality, we get

$$\begin{bmatrix} -A^{-1}B \\ I \end{bmatrix}^* (Q + \epsilon I) \begin{bmatrix} -A^{-1}B \\ I \end{bmatrix} \preceq 0$$

for some scalar $\epsilon > 0$. Hence, there exists a diagonal $P \succeq 0$ such that

$$Q + \epsilon I + \begin{bmatrix} A^T P + PA & PB \\ B^T P & 0 \end{bmatrix} \preceq 0$$

Adding a small multiple of the identity to P gives $P \succ 0$ such that

$$Q + \begin{bmatrix} A^T P + PA & PB \\ B^T P & 0 \end{bmatrix} \prec 0$$

so also (14.3) holds with strict inequality. Hence the proof is complete. \square

An analogous discrete time result is stated here and proved in the appendix:

Theorem 15: Let $A \in \mathbb{R}_+^{n \times n}$ be Schur, while $B \in \mathbb{R}_+^{n \times m}$ and the pair (A, B) is controllable. Suppose that all entries of $Q \in \mathbb{R}^{(n+m) \times (n+m)}$ are nonnegative, except for the last m diagonal elements. Then the following statements are equivalent:

(15.1) For $\omega \in [0, \infty]$ is true that

$$\begin{bmatrix} (e^{i\omega} I - A)^{-1} B \\ I \end{bmatrix}^* Q \begin{bmatrix} (e^{i\omega} I - A)^{-1} B \\ I \end{bmatrix} \preceq 0$$

(15.2) $\begin{bmatrix} (I - A)^{-1} B \\ I \end{bmatrix}^* Q \begin{bmatrix} (I - A)^{-1} B \\ I \end{bmatrix} \preceq 0$.

(15.3) There exists a diagonal $P \succeq 0$ such that

$$Q + \begin{bmatrix} A^T P A - P & A^T P B \\ B^T P A & B^T P B \end{bmatrix} \preceq 0$$

Moreover, if all inequalities are taken to be strict, then the equivalences hold even without the controllability assumption.

X. CONCLUSIONS

The results above demonstrate that the monotonicity properties of positive systems and positively dominated systems bring remarkable benefits to control theory. Most important is the opportunity for scalable verification and synthesis of distributed control systems with optimal input-output performance. In particular, the linear programming solution comes with certificates that enable distributed and scalable verification of global optimality, without access to a global model anywhere.

Many important problems remain open for future research. Here are two examples:

- How can the scalable methods for verification and synthesis be extended to monotone nonlinear systems?
- How can the optimization of static controller gains in this paper be extended to scalable methods for optimization of dynamic controllers?

XI. ACKNOWLEDGMENTS

The author is grateful for suggestions and comments by numerous colleagues. In particular, Theorem 5 was updated based on suggestions by Andrej Ghulchak. The work has been supported by the Swedish Research Council through the Linnaeus Center LCCC.

REFERENCES

- [1] David Angeli and Eduardo D. Sontag. Monotone control systems. *IEEE Transactions on Automatic Control*, 48(10):1684–1698, October 2003.
- [2] Paul C. Bell, Jean-Charles Delvenne, Raphaël M. Jungers, and Vincent D. Blondel. The continuous Skolem-Pisot problem. *Theoretical Computer Science*, 411:3625–3634, 2010.
- [3] Abraham Berman and Robert J. Plemmons. *Nonnegative matrices in the mathematical sciences*. Classics in Applied Mathematics. SIAM, 1994.
- [4] G. Birkhoff. Extensions of Jentzsch's theorem. *Trans. Amer. Math. Soc.*, 85:219–227, 1957.
- [5] Vincent D. Blondel and Natacha Portier. The presence of a zero in an integer linear recurrent sequence is NP-hard to decide. *Linear Algebra and its Applications*, 351–352:91–98, 2002.
- [6] Stephen Boyd and Lieven Vandenbergh. *Convex Optimization*. Cambridge University Press, 2004.
- [7] Lorenzo Farina and Sergio Rinaldi. *Positive Linear Systems: Theory and Applications*. Wiley, 2000.
- [8] M. W. Hirsch. Differential equations and convergence almost everywhere in strongly monotone flows. In *Contemporary Mathematics*, volume 17, pages 267–285. Amer. Math. Soc., Providence, RI, 1983.

- [9] M.W. Hirsch and Hal Smith. Monotone dynamical systems. In *Handbook of Differential Equations: Ordinary Differential Equations*, volume 2, chapter 4, pages 239–357. Elsevier, 2006.
- [10] L. Hörmander. *The analysis of linear partial differential operators I*. Springer Verlag, 1985.
- [11] John A. Jacquez. *Compartmental Analysis in Biology and Medicine*. BioMedware, 3rd edition, 1996.
- [12] Tadeusz Kaczorek. *Positive 1D and 2D systems*. Springer, 2002.
- [13] Sunyoung Kim and Masakazu Kojima. Exact solutions of some non-convex quadratic optimization problems via SDP and SOCP relaxations. *Computational Optimization and Applications*, 26:143–154, 2003.
- [14] Elon Kohlberg and John W. Pratt. The contraction mapping approach to the perron-frobenius theory: Why Hilbert's metric? *Mathematics of Operations Research*, 7(3):198–210, May 1982.
- [15] Javad Lavaei, Anders Rantzer, and Steven Low. Power flow optimization using positive quadratic programming. In *18th IFAC World Congress*, Milano, Italy, August 2011.
- [16] Patrick De Leenheer and Dirk Aeyels. Stabilization of positive linear systems. *Systems & Control Letters*, 44(4):259–271, November 2001.
- [17] Wassily W. Leontief, editor. *Input-Output Economics*. Oxford University Press, New York, 2nd edition, 1986.
- [18] David G. Luenberger. *Introduction to dynamic systems: Theory, models, and applications*. Wiley (New York), 1979.
- [19] A. Megretski, 2010. Personal Communication.
- [20] Alexander Megretski and Anders Rantzer. System analysis via Integral Quadratic Constraints. *IEEE Transactions on Automatic Control*, 42(6):819–830, June 1997.
- [21] Luc Moreau. Stability of continuous-time distributed consensus algorithms. In *Proceedings of 43rd IEEE Conference on Decision and Control*, Bahamas, 2004.
- [22] Ait Rami. Solvability of static output-feedback stabilization for LTI positive systems. *Systems & Control Letters*, 60:704–708, 2011.
- [23] M. Ait Rami and F. Tadeo. Controller synthesis for positive linear systems with bounded controls. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 54(2):151–155, 2007.
- [24] Anders Rantzer. Distributed control of positive systems. In *50th IEEE Conference on Decision and Control and European Control Conference*, Orlando, Florida, USA, December 2011. Full paper available from arXiv.org/abs/1203.0047.
- [25] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton Landmarks in Mathematics and Physics. Princeton University Press, 1997.
- [26] E. Seneta. *Non-negative Matrices and Markov Chains*. Springer Series in Statistics. Springer, New York, 2nd edition, 1981.
- [27] Rodolphe Sepulchre. Consensus on nonlinear spaces. In *Proceedings of the 8th IFAC Symposium on Nonlinear Control Systems*, Sep 2010.
- [28] Takashi Tanaka and Cédric Langbort. The bounded real lemma for internally positive systems and H-infinity structured static state feedback. *IEEE Transactions on Automatic Control*, 56(9):2218–2223, September 2011.
- [29] J.N. Tsitsiklis. *Problems in Decentralized Decision Making and Computation*. PhD thesis, MIT, 1984.
- [30] J.C. Willems. Lyapunov functions for diagonally dominant systems. *Automatica*, pages 519–523, 1976.

XII. APPENDIX

Proof of Theorem 4 First note that the maximum $\max_{\omega} |\mathbf{G}(e^{i\omega})|$ must be attained at $\omega = 0$ since

$$\begin{aligned} |\mathbf{G}(e^{i\omega})| &= \left| D + e^{-i\omega} \sum_{t=0}^{\infty} C e^{-it\omega} A^t B \right| \\ &\leq D + \sum_{t=0}^{\infty} C A^t B = D + C(I - A)^{-1} B = \mathbf{G}(1) \end{aligned}$$

Hence $\|\mathbf{G}\|_{\infty} < \gamma$ may equivalently be written

$$C(I - A)^{-1} B + D < \gamma$$

Suppose that (4.1) holds. Then $C(I - A)^{-1} B + D < \gamma$. By Proposition 2 there exists $\xi > 0$ such that $A\xi < \xi$. Define

$x = \xi + (I - A)^{-1} B$. Then $x > 0$ since $(I - A)^{-1} \geq 0$. Moreover

$$x - Ax = \xi - A\xi + B > B$$

If ξ is sufficiently small, we also get $Cx + D < \gamma$ so

$$\begin{bmatrix} A & B \\ \gamma^{-1}C & \gamma^{-1}D \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} < \begin{bmatrix} x \\ 1 \end{bmatrix} \quad (17)$$

and (4.2) holds. Conversely, (4.2) implies (17) for some $x > 0$, so A is Schur and

$$(I - A)^{-1} B < x \quad C(I - A)^{-1} B + D < Cx + D < \gamma$$

Hence (4.1) follows and equivalence between (4.1) and (4.2) has been established.

Next assume that (4.3) holds. For $w = 0$, the inequality implies $|x(t+1)|_P^2 \leq |x(t)|_P^2$ with strict inequality for $x(t) \neq 0$, so A must be Schur. Summing the inequality in (4.4) along solutions to $x(t+1) = Ax(t) + Bw(t)$ with $x(0) = 0$ gives

$$\sum_{t=0}^T |Cx(t) + Dw(t)|^2 dt \leq \gamma^2 \sum_{t=0}^T |w(t)|^2 dt$$

so $\|\mathbf{G}\|_{\infty} \leq \gamma$ and (4.1) follows. For the opposite implication, assume that (4.2) holds. Then Proposition 2 proves existence of a diagonal $P \succ 0$ such that

$$\begin{bmatrix} A & B \\ \gamma^{-1}C & \gamma^{-1}D \end{bmatrix}^T \begin{bmatrix} P & 0 \\ 0 & \gamma^2 \end{bmatrix} \begin{bmatrix} A & B \\ \gamma^{-1}C & \gamma^{-1}D \end{bmatrix} \prec \begin{bmatrix} P & 0 \\ 0 & \gamma^2 \end{bmatrix}$$

Multiplying with $[x^T \ w^T]$ from the left and its transpose from the right gives

$$(Ax + Bw)^T P (Ax + Bw) + (Cx + Dw)^2 \leq x^T P x + \gamma^2 w^2$$

with strict inequality for $(x, w) \neq (0, 0)$. Condition (4.3) follows and the equivalence between (4.1) \Leftrightarrow (4.2) \Leftrightarrow (4.3) has been established.

Assuming that (4.4) holds, consider $x(t) > 0$, $w(t) > 0$ and rewrite the inequality as

$$p^T (Ax + Bw) + Cx + Dw < p^T x + \gamma w$$

Hence

$$\begin{bmatrix} A & B \\ \gamma^{-1}C & \gamma^{-1}D \end{bmatrix}^T \begin{bmatrix} p/\gamma \\ 1 \end{bmatrix} < \begin{bmatrix} p/\gamma \\ 1 \end{bmatrix} \quad (18)$$

and (4.2) follows by Proposition 2.

Finally, assuming (4.2), Proposition 2 gives (18) for some $p > 0$, so

$$\begin{aligned} &p^T |x(t+1)| + |Cx(t) + Dw(t)| \\ &\leq p^T (A|x(t)| + B|w(t)|) + C|x(t)| + D|w(t)| \\ &\leq p^T |x(t)| + \gamma |w(t)| \end{aligned}$$

with strict inequality when $(x(t), w(t)) \neq (0, 0)$. This proves (4.4) and the proof is complete. \square

Proof of Theorem 15. The theorem can be proved in analogy with the proof of Theorem 14. Alternatively, it can be derived from Theorem 14 using a bilinear transformation in the following way:

Instead of $e^{i\omega}$, one can parametrize the unit circle as $\frac{1+i\omega}{1-i\omega}$. Hence (15.1) is equivalent to saying that

$$\begin{bmatrix} x \\ w \end{bmatrix}^* Q \begin{bmatrix} x \\ w \end{bmatrix} \leq 0$$

for all solutions (ω, x, w) to the equation

$$\left(\frac{1+i\omega}{1-i\omega} I - A \right) x = Bw$$

Alternatively, introducing

$$\begin{aligned} \hat{A} &= (A - I)(A + I)^{-1} \\ \hat{B} &= 2(A + I)^{-1}B \\ \hat{x} &= x + Ax + Bw \\ S &= \begin{bmatrix} (A + I)^{-1} & -(A + I)^{-1}B \\ 0 & I \end{bmatrix} \\ \hat{Q} &= S^T Q S \end{aligned}$$

the condition can be re-written as the statement that

$$\begin{bmatrix} \hat{x} \\ w \end{bmatrix}^* \hat{Q} \begin{bmatrix} \hat{x} \\ w \end{bmatrix} \leq 0$$

for all solutions $(\omega, \hat{x}, \hat{w})$ to the equation $(i\omega I - \hat{A})\hat{x} = \hat{B}w$. According to Theorem 14, this is equivalent to existence of a diagonal $P \succeq 0$ such that

$$\hat{Q} + \begin{bmatrix} \hat{A}^T P + P \hat{A} & P \hat{B} \\ \hat{B}^T P & 0 \end{bmatrix} \preceq 0$$

Multiplying by $\begin{bmatrix} A + I & B \\ 0 & I \end{bmatrix}$ from the right and its transpose from the left, the matrix inequality (after trivial manipulations) becomes

$$Q + 2 \begin{bmatrix} A^T P A - P & A^T P B \\ B^T P A & B^T P B \end{bmatrix} \preceq 0$$

Replacing $2P$ by P makes the proof complete. \square